2К

УДК 004.8

Ю.В. Крак, О.В. Бармак, Г.М. Єфімов
Київський національний університет імені Тараса Шевченка, Україна
Хмельницький національний університет, Україна
Таврійський державний агротехнологічний університет, Україна

Інформаційна технологія розпізнавання емоційної міміки на обличчі людини

В роботі наведено цілісну інформаційну технологію аналізу мімічних виразів емоцій на обличчі людини. Міміка емоційних станів людини описана у вигляді лінійної комбінації деяких базових станів. Параметри моделі отримуються за допомогою методу гнучких шаблонів з використанням NURBS-кривих. Довільна емоція конкретної людини розпізнається як опукла комбінація станів запропонованого базового простору.

В роботі розглядається задача моделювання мімічних виразів емоцій згідно із запропонованим в [1] формальним описом базових емоцій. Для пошуку простору характеристичних ознак, побудови базису цього простору, відтворення похідних емоційних станів з наступним застосуванням опуклої комбінації, пропонується наступне [2], [3]:

– створення множини фотографічних зображень мімічних проявів на обличчі, які відповідають ситуаціям, у яких виникають базові емоції, опис міміки, властивої цим емоціям;

– аналіз отриманої множини з метою виявлення областей, які містять характеристичні ознаки емоцій та їх опис;

– створення у просторі характеристичних ознак базису для наступного розкладу по ньому довільних векторів мімічних проявів емоційних станів (як опуклої комбінації базисних емоційних станів).

В [2] запропоновано для моделювання міміки емоційних станів людини використовувати лінійну опуклу комбінацію деяких базових станів. В ході цих досліджень було отримано 21 характеристичну ознаку, комбінація яких утворює базис мімічних виразів емоцій. Тобто мімічні вирази емоції (Em) були представлені у вигляді вектора:

$$Em_i^{\eta} = (\mu_1, \cdots, \mu_{21}), i = 1 \cdots 8, \qquad (1)$$

де $\mu_j \in [0;1]$ – характеристична мімічна ознака (при $\mu = 0$ – немає ознаки, а при $\mu = 1$ вплив ознаки максимальний); перелік ознак наведений в табл. 1.

Набір 8 векторів утворює базис (B_{ij}, де i = 1, …, 21, j = 1, …, 8) простору мімічних проявів емоційних станів.

Таким чином, довільний вектор $b=(\mu 1, \ldots, \mu 21)$, отриманий шляхом аналізу зображення з якимось емоційним станом, можна розкласти по базису B і отримати опис емоції як опуклої комбінації:

$$x = (B^T B)^{-1} B^T b, \qquad (2)$$

за умови

$$\det((B^T B)^{-1}) > 0, \qquad (3)$$

де B – базисна матриця емоційних станів (табл. 1);

B^T – транспонована матриця B;

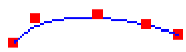
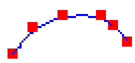
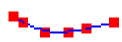
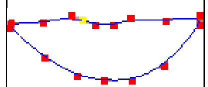
b – вектор, який описує мімічний прояв довільного емоційного стану;

$x = (\alpha_1,\ldots,\alpha_8)$, де α_i – коефіцієнти опуклої комбінації ($\Sigma\alpha_i=1$, $\alpha_i \in [0;1]$) для кожної з 8 базових емоцій.

У запропонованій в [2] моделі базис простору мімічних ознак емоційних станів будується на основі апріорного досвіду експериментатора, вимагає певної кваліфікації і, відповідно, дає неоднозначний результат – констатація одного і того ж емоційного м'язового прояву у різних людей розрізняється. Для того щоб перейти від феноменологічного визначення характеристичних мімічних ознак до певної їх формалізації, у [3] запропоновано використати власну модифікацію методу моделей, які деформуються. Як параметричні криві, для виділення характерних ознак рис обличчя, запропоновано використати моделі, які задаються з допомогою нерівномірних раціональних базисних сплайнів – NURBS-кривих [4], [5].

В рамках проведених досліджень була побудована модель, яка складається з гнучких шаблонів – NURBS-кривих. Виходячи з того, що опорні точки NURBS-кривої однозначно визначають саму криву, до розгляду брались тільки вектори опорних точок. Для побудови базису були використані наступні вектори опорних точок (табл. 1).

Таблиця 1 – Мімічні прояви для формування базових емоцій

ознака	Опис мімічних проявів в розрізі областей обличчя			Для брів, очей та рота – вектори опорних точок p_i^j	
	Область обличчя		Мімічний прояв		
μ_1	Область чола і брів	Чоло	Зморшки в центрі чола	Для зморшок – $\mu_j \in [0;1]$ (при $\mu = 0$	
μ_2			Одна горизонтальна зморшка	немає зморшки, а при $\mu = 1$	
μ_3			Між бровами вертикальна зморшка	зморшка максимальна)	
μ_4		Брови	Внутрішні кутики підняті вверх	$p_i^1, i = 0 \cdots 4$	 Ліва брова
μ_5			Опущені та зведені		
μ_6			Припідняті		
μ_7			Підняті та зведені		
μ_8	Область очей (очі, повіки, основа носа)	Верхні повіки	Внутрішні кутики підняті	$p_i^2, i = 0 \cdots 5$	 Ліва верхня повіка
μ_9			Напружені		
μ_{10}			Підняті (видно склеру)		
μ_{11}			Припідняті		
μ_{12}		Нижні повіки	Припідняті та ненапружені	$p_i^3, i = 0 \cdots 5$	 Ліва нижня повіка
μ_{13}			Припідняті та напружені		
μ_{14}		Зморшки	«Гусячі лапки» біля зовнішніх кутиків		
μ_{15}			Зморшка під повіками		
μ_{16}	Нижня частина обличчя (ніс, щоки, рот)	Рот	Закритий, губи стиснуті	$p_i^4, i = 0 \cdots 16$	 Губи
μ_{17}			Розкритий		
μ_{18}		Губи (лінія, кутики)	Кутики губ відтягнуті в сторони та припідняті		
μ_{19}			Розтягнуті та напружені		
μ_{20}			Кутики губ опущені		
μ_{21}		Зморшки	Зморшка від носа до кутиків губ		

У [2], [3] запропонована математична модель та цілісна інформаційна технологія для автоматичного визначення довільного емоційного стану конкретної людини як опуклої комбінації деяких базових станів. Для цього з допомогою математичної моделі та оригінального програмного забезпечення створюється базовий простір емоційних станів конкретної людини. В подальшому довільний емоційний прояв цієї людини розкладається як опукла комбінація емоційних станів в цьому просторі. Але в цій моделі гнучкі шаблони налаштовувалися на прояв конкретної міміки з допомогою ручного редагування певної кількості параметрів – контрольних точок NURBS-кривих на поверхні зображення. Подальші дослідження у цьому напрямку були направлені на модифікацію методу контурних моделей, які деформуються. Для цілей автоматичного деформування моделей було запропоновано використовувати апроксимацію *B*-сплайн кривих [4].

1 B-сплайн апроксимація кривих

Скористаємося простою властивістю NURBS-кривих: при $w_i = 1$ NURBS-крива зводиться до *B*-сплайн кривої. Ця проста властивість випливає з ідентичності опорних точок (p_i) в однорідній формі і рівності знаменника 1. Беручи до уваги, що при моделюванні гнучких шаблонів приймалося $w_i = 1$, то для спрощення апроксимації можна перейти до *B*-сплайн кривих.

Задача *B*-сплайн апроксимації є задачею підгонки *B*-сплайн кривої з *K* опорними точками $p = [p_0, \cdots, p_{K-1}]^T$ до точкової кривої $d = [d_0, \cdots, d_{M-1}]^T$, де $M > K$ (зазвичай $M >> K$) для значень параметра $u_0, \cdots, u_{M-1}$. Така задача апроксимації приводить до перевизначеної системи лінійних рівнянь $N \cdot p = d$:

$$\begin{bmatrix} N_0(u_0) & \cdots & N_{K-1}(u_0) \\ N_0(u_1) & \cdots & N_{K-1}(u_1) \\ \vdots & \ddots & \vdots \\ N_0(u_{M-1}) & \cdots & N_{K-1}(u_{M-1}) \end{bmatrix} \cdot \begin{bmatrix} p_0 \\ \vdots \\ p_{K-1} \end{bmatrix} = \begin{bmatrix} d_0 \\ d_1 \\ \vdots \\ d_{M-1} \end{bmatrix}, \qquad (4)$$

де $N_i(u)$ – *B*-сплайн базисна функція.

Одним з шляхів отримання рішення перевизначеної системи лінійних рівнянь (4) є:

$$N^T N \cdot p = N^T \cdot d , \qquad (5)$$

звідки

$$p = \left(N^T N\right)^{-1} \cdot N^T d , \qquad (6)$$

при

$$\det\left(\left(N^T N\right)^{-1}\right) > 0 . \qquad (7)$$

Для застосування *B*-сплайн апроксимації потрібно вміти отримувати на зображенні точкові криві $d = [d_0, \cdots, d_{M-1}]^T$, які відповідають потрібним контурам, з тим, щоб далі застосувати перетворення (5 – 7).

2 Технології оконтурення та скелетизації зображень

Існує багато технологій для отримання на зображенні точкових кривих $d = [d_0, \cdots, d_{M-1}]^T$, які відповідають контурам брів, очей та рота. В основному вони базуються на отриманні контуру зображення як різкої границі між елементами зображення (з допомогою згорток, аналізу кольору тощо) з наступною скелетизацією (отриманням контуру одиничної товщини).

В даній роботі для оконтурення зображення пропонується застосувати імітацію роботи зорових рецепторів ока людини. Відомо [6], що очне яблуко знаходиться в неперервному мікрорусі. Питання про ці мікрорухи має неоднозначне трактування. Можна передбачити, що ці мікрорухи є необхідною умовою функціонування апарату виділення контурів на зображенні. Для перевірки цього заставимо рецептори сітківки штучного ока зафіксувати запропоноване зображення, потім незначним чином (наприклад, на 1 точку) здвинемо зображення в сторону і знову дамо можливість рецепторам ока зафіксувати його. В цей момент на виходах рецепторів з'явиться відносна зміна сигналу. Візьмемо значення змін рецепторів та занесемо їх у відповідні точки на зображенні – отримаємо контури зображення (рис. 1).



Рисунок 1 – Зображення до та після оконтурення

Імітація вмісту рецепторів на сітківці ока проходить наступним чином. Є зображення та напрямок мікроруху (наприклад, діагонально на M точок). Спочатку конкретний рецептор «бачить» точку з координатами (x, y), а після мікроруху – з координатами (x-M, y-M). Різниця кольорових площин між вхідною точкою та точкою, яка виявилась на її місці в результаті мікроруху, – це і є відносна зміна вхідного дратівного сигналу (для конкретного рецептора).

Отримані таким чином контури потрібно привести до «скелетного» вигляду. Тобто потрібно виділити деяку середню лінію, яка б правильно відображала структуру контуру. Для цього застосуємо відомий алгоритм Зонга – Суня [7]. Отриманий таким чином «скелет» контуру губ зображено на рис. 2.

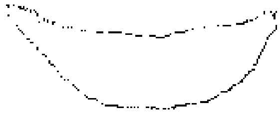


Рисунок 2 – Точкова крива губ, отримана після скелетизації контуру

Основна ідея алгоритму Зонга – Суня полягає в тому, що на кожному кроці, проходячи по зображенню рамкою 3 × 3, перевіряється приналежність кожного пікселя до границі заданої зв'язної області. Якщо умови перевірки виконуються, то піксель вилучається з області. Незважаючи на кількість виконаних кроків, область залишиться зв'язаною, у крайньому випадку вона виродиться в лінію товщиною в один піксель.

3 Технологія аналізу мімічних виразів емоцій

Для аналізу мімічних виразів емоцій запропонована наступна цілісна інформаційна технологія [2], [3]:

1) для конкретної людини отримуємо множину із восьми фотографічних зображень обличчя, на якому відтворена мімічна реакція на ситуації, які відповідають базовим емоціям: радість, горе, надія, страх, задоволення, гнів, інтерес, зневага;

2) у відповідному програмному забезпеченні нормуємо фотографічні зображення (по відстані між центрами очей);

3) за допомогою відомих технологій оконтурення та скелетування зображень (п. 2) отримуємо контури наступних частин обличчя: зморшки в області чола, брів, верхніх повік, нижніх повік, зморшки «гусячі лапки» біля зовнішніх кутиків очей, зморшки під повіками, рот, зморшки від носа до кутиків губ;

4) використовуючи гнучкі шаблони у вигляді NURBS-кривих [3] та *B*-сплайн апроксимацію (п. 1), отримаємо множину опорних точок NURBS-кривої для шаблонів, які характеризують кожну з восьми базових емоцій (табл. 1):

$p_i^{1,(e)} = \left[x_i^{1,(e)}, y_i^{1,(e)}\right]^T, i = 0\ldots4, e = 1\ldots8$ – шаблон лівої брови для 8 емоційних станів;

$p_i^{2,(e)} = \left[x_i^{2,(e)}, y_i^{2,(e)}\right]^T, i = 0\ldots5, e = 1\ldots8$ – шаблон лівої верхньої повіки для 8 емоційних станів;

$p_i^{3,(e)} = \left[x_i^{3,(e)}, y_i^{3,(e)}\right]^T, i = 0\ldots5, e = 1\ldots8$ – шаблон лівої нижньої повіки для 8 емоційних станів;

$p_i^{4,(e)} = \left[x_i^{4,(e)}, y_i^{4,(e)}\right]^T, i = 0\ldots16, e = 1\ldots8$ – шаблон губів для 8 емоційних станів;

та для опису зморшок отримаємо деякі характеристичні мімічні ознаки (табл. 1):

$\mu^{(e)} = \left[\mu_1^{(e)}, \mu_2^{(e)}, \mu_3^{(e)}, \mu_{14}^{(e)}, \mu_{15}^{(e)}, \mu_{21}^{(e)}\right]^T, e = 1\ldots8$ – характеристичні мімічні ознаки (табл. 1);

5) з восьми отриманих множин опорних точок NURBS-кривих (шаблонів) та вектора характеристичних мімічних ознак для зморшок побудуємо базис емоційних станів (B) конкретної людини:

$$B = \begin{bmatrix} \mu^{(1)} & \cdots & \mu^{(8)} \\ \mathrm{P}^{1,(1)} & \cdots & \mathrm{P}^{1,(8)} \\ \mathrm{P}^{2,(1)} & & \mathrm{P}^{2,(8)} \\ \mathrm{P}^{3,(1)} & & \mathrm{P}^{3,(8)} \\ \mathrm{P}^{4,(1)} & & \mathrm{P}^{4,(8)} \end{bmatrix}, \quad \text{де} \quad \mathrm{P}^{j,(k)} = \begin{bmatrix} p_0^{j,(k)} \\ \vdots \\ p_{n_j}^{j,(k)} \end{bmatrix}, \quad \text{де} \quad \begin{array}{l} j = 1\ldots4 \\ k = 1\ldots8 \\ n_1 = 4, n_2, n_3 = 5, n_4 = 16. \end{array} \tag{8}$$

В подальшому для аналізу довільного зображення обличчя цієї людини повторимо пункти 1) – 4) для зображення довільної емоції та отримаємо вектор:

$$b = \begin{bmatrix} \mu^{(*)} \\ \mathrm{P}^{1,(*)} \\ \mathrm{P}^{2,(*)} \\ \mathrm{P}^{3,(*)} \\ \mathrm{P}^{4,(*)} \end{bmatrix}, \quad \text{де} \quad \mathrm{P}^{j,(*)} = \begin{bmatrix} p_0^{j,(*)} \\ \vdots \\ p_{n_j}^{j,(*)} \end{bmatrix}, \text{де} \quad \begin{array}{l} j = 1\ldots4 \\ \\ n_1 = 4, n_2, n_3 = 5, n_4 = 16. \end{array} \tag{9}$$

Розкладемо його по побудованому базису *B* (8):

$$B^T B \cdot x = B^T \cdot b, \tag{10}$$

звідки

$$x = \left(B^T B\right)^{-1} \cdot B^T b\,, \tag{11}$$

при

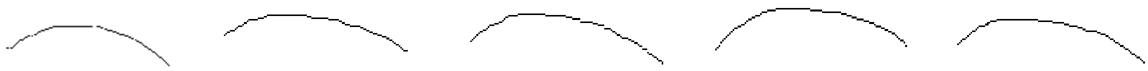
$$\det\left(\left(B^T B\right)^{-1}\right) > 0\,. \tag{12}$$

Коефіцієнти розкладу $x = (\alpha_1, \ldots, \alpha_8)$ будуть вказувати на конкретний вклад кожної з восьми базових емоцій у довільну емоцію b.

4 Результати дослідження

За умови переходу від феноменологічного опису емоцій до ситуацій, у яких емоції виникають, була створена множина фотографічних зображень базових емоцій ([1], п. 1).

Далі з отриманих зображень виділялися описаними вище способами (п. 2) необхідні для подальшої обробки контури (брови, очі, губи тощо). Рис. 3 містить контури правої брови для емоцій: радість, горе, надія, страх, задоволення.



Радість Горе Надія Страх Задоволення

Рисунок 3 – Контури правої брови для деяких емоцій

На рис. 4 контури правої брови представлені у вигляді точкової кривої.

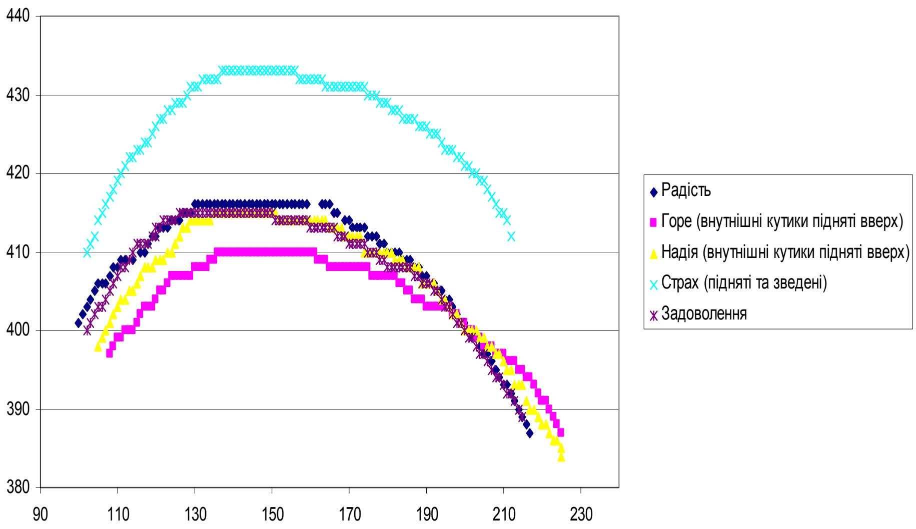


Рисунок 4 – Графік точкових кривих правої брови для деяких емоцій

На графіку видно, що положення отриманих контурів правої брови відповідають опису міміки емоцій. Тобто видно, що для емоцій задоволення та радості особливого прояву міміки немає, для емоцій горя та надії – внутрішні кутики підняті догори, а для емоції страху – брова піднята та зведена.

До отриманих контурів застосовувались перетворення (5 – 7) для отримання множин керуючих точок NURBS-кривих. Рис. 5 містить графік контуру та відповідну цьому контуру NURBS-криву для положення правої брови при емоції радості.

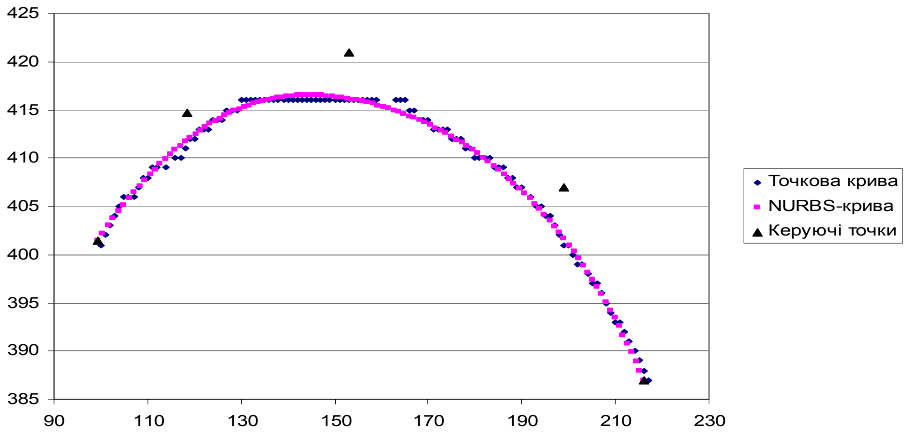


Рисунок 5

Рисунок 6 містить результат прив'язки гнучких шаблонів (NURBS-кривих) до відповідних фрагментів обличчя для всіх базових емоцій.



Рисунок 6 – Представлення базових емоцій у контурному вигляді з допомогою NURBS-кривих для конкретного актора

З восьми отриманих множин опорних точок NURBS-кривих (шаблонів) та вектора характеристичних мімічних ознак для зморшок був побудований базис (8) емоційних станів (B) конкретної людини.

Аналогічним чином з фотографічного зображення довільної емоції, що відтворює ситуацію, у якій виникає відчуття провини (рис. 7), був побудований відповідний вектор b (9).

Із застосуванням перетворення (10 – 12) був проведений розклад отриманого вектора b по базису B.

Для емоційного стану «провина» були отримані наступні коефіцієнти опуклої комбінації базових станів:

$$\alpha_1 = 0, \alpha_2 = 0.7, \alpha_3 = 0, \alpha_4 = 0, \alpha_5 = 0.3, \alpha_6 = 0, \alpha_7 = 0, \alpha_8 = 0, \sum_{i=1}^{8} \alpha_i = 1,$$

де α_i відповідають наступним емоційним станам: радість, горе, надія, страх, задоволення, гнів, інтерес та зневага. Згідно з [1], стан, який складається з комбінацій задоволення ($\alpha_5 = 0.3$) та горя ($\alpha_2 = 0.7$), відповідає емоції «провина».

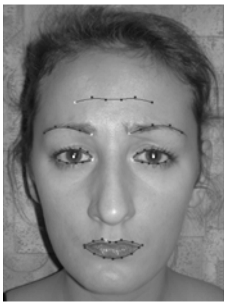


Рисунок 7 – Відтворення ситуації, у якій виникає емоція провини

Висновки

Запропонована математична модель та цілісна інформаційна технологія для автоматичного визначення довільного емоційного стану конкретної людини як опуклої комбінації деяких базових станів. Для цього з допомогою математичної моделі та оригінального програмного забезпечення створюється базовий простір емоційних станів конкретної людини. В подальшому довільний емоційний прояв цієї людини розкладається як опукла комбінація емоційних станів в цьому просторі.

Для побудови базису простору емоційних станів використовуються гнучкі шаблони контурів основних зон обличчя. Гнучкі шаблони описуються за допомогою NURBS-кривих. Налаштування шаблону на точковий контур конкретного зображення проходить з допомогою B-сплайн апроксимації, шляхом розв'язку перевизначеної неоднорідної системи лінійних рівнянь.

Запропонована технологія має практичну цінність в системах візуального контролю за операторами складних виробництв (атомна енергетика тощо) для автоматичного контролю за їх емоційним станом.

Література

1. Леонтьев В.О. Классификация эмоций. – Одесса: Инновационно-ипотечный центр, 2002. – 84 с.
2. Крак Ю.В., Бармак О.В., Єфімов Г.М. Синтез мімічних виразів емоцій на основі формальної моделі // Штучний інтелект. – 2007. – № 2. – С. 22-31.
3. Крак Ю.В., Бармак О.В., Єфімов Г.М. Використання контурних моделей для побудови базису простору мімічних виразів емоцій // Штучний інтелект. – 2007. – № 4. – С. 288-296.
4. Piegl, Les and Tiller, Wayne. The NURBS Book, 2nd Edition. – Berlin, Germany: Springer-Verlag, 1996.
5. C. deBoor. A Practical Guide to Splines. – New York: Springer-Verlag, 1978.
6. Демидов В.Е. Как мы видим то, что видим. – М.: Знание, 1987. – 240 с.
7. Zhang T.Y., Suen C.Y. A fast parallel algorithm for thinning digital patterns // Commun. ACM. – 1984. – Vol. 27, № 3. – P. 236-239.

Ю.В. Крак, О.В. Бармак, Г.М. Ефимов

Информационная технология распознавания эмоциональной мимики на лице человека

В работе приведена целостная информационная технология анализа мимических выражений эмоций на лице человека. Мимика эмоциональных состояний человека описана в виде линейной комбинации некоторых базовых состояний. Параметры модели получают при помощи метода гибких шаблонов с использованием NURBS-кривых. Произвольная эмоция конкретного человека распознается как выпуклая комбинация состояний предложенного базового пространства.

Стаття надійшла до редакції 04.02.2008.